Zeszyty Naukowe Wydziału Elektrotechniki i Automatyki Politechniki Gdańskiej Nr 51

XXVI Seminarium

ZASTOSOWANIE KOMPUTERÓW W NAUCE I TECHNICE 2016

Oddział Gdański PTETiS

ANALYSIS OF EFFECTIVENESS AND COMPUTATIONAL COMPLEXITY OF TREND REMOVAL METHODS

Łukasz LENTKA, Janusz SMULKO

Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics
Department of Metrology and Optoelectronics
tel.:+48 58 348 6095 e-mail: lukasz.lentka@pg.gda.pl

Abstract: The paper presents a method of processing measurement data due to remove slowly varying component of the trend occurring in the recorded waveforms. Comparison of computational complexity and trend removal efficiency between some commonly used methods is presented. The impact of these procedures on probability distribution and power spectral density is shown. Effectiveness and computational complexity of these methods depend essentially on nature of the removed trend. This paper describes several procedures: Moving Average Removal (MAR), fitting a polynomial of degree appropriate to the analyzed data, Empirical Mode Decomposition (EMD).

Keywords: trend removal, high-pass filtering, Empirical Mode Decomposition (EMD).

1. INTRODUCTION

Recorded time series, especially noise records, exhibit often slowly changing component – trend. Trend removal (also called drift removal) is a data processing procedure which has many practical applications [1-6]. In general, trend removal can be treated as a high-pass filtration of recorded data. In order to remove the trend component, we can use a high-pass filter (analog or digital) at the stage of data acquisition. However, we often have to deal with the data, in which a trend component is needed to carry out other analysis and cannot be removed at the stage of data registration. Furthermore, signal may be non-stationary and high-pass filtering will not be efficient. In such cases, we should consider other, often more advanced methods [7-10].

Chapter 2 presents several trend removal methods, classical like high-pass filtering, Moving Average Removal (MAR) and polynomial fitting, as well as more sophisticated method Empirical Mode Decomposition (EMD). In chapter 3 we present simulation results, which consider the effect of trend removing on the probability distribution and power spectral density. Chapter 4 concludes the paper.

2. TREND REMOVAL METHODS

2.1. High-pass filtering

The most general method of trend removal is high-pass filtering. We can say that all trend removal methods, even very complex, can be classified as high-pass filtering. Figure 1 presents idea of trend removal. On the left side there is graph of noise imposed on the extremely low frequency trend, which is the input data. On the right side we see noise with the removed trend after high-pass filter (HPF) and noise after low-pass filter (LPF). The red rectangle represents an operation of trend removal.

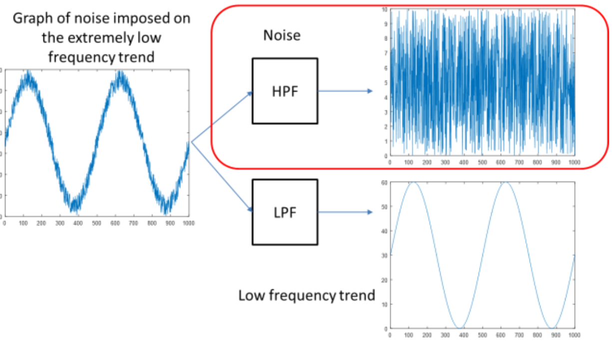

Fig. 1. General idea of trend removal operation

However, high-pass filtering at signal acquisition stage is rather a simply method, it losses completely information about trend component, which can be used for further analysis.

2.2. Moving Average Removal

Well-known Moving Average (MA) method is a low-pass filtering, but Moving Average Removal (MAR) is a high-pass filtering [1]. Principles of operation of the MAR can be illustrated by following equations:

$$m_n = \frac{1}{2p+1}(x_{n-p} + \ldots + x_{n-1} + x_n + x_{n+1} + \ldots + x_{n+p}) \quad (1)$$

$$y_n = x_n - m_n \quad (2)$$

where: n – sample number, p – filter parameter (averaging over $2p+1$ points), x_n – n-th input sample (raw data), m_n – average of the segment from point n-p to n+p, y_n – n-th output sample (processed data).

The p parameter have big influence on resulting MAR filter as can be seen by analysis of its transfer function $H(f)$ given by the formula:

$$H(f) = 1 - \frac{1}{2p+1}\frac{\sin[(2p+1)\pi f / f_s]}{\sin[\pi f / f_s]} \quad (3)$$

where: f – frequency, f_s – sampling frequency.

The MAR filtering procedure does not introduce a phase shift and effectively removes drift component, but also a large part of low-frequency components of the input signal [1]. It is clear from the equation (3) that the passband of the filter is within the frequency range from $f_s/(2(2p+1)$ to $f_s/2$. Consequently, if p is small the range of the filter is quite narrow. On the other hand, for a high order filter (when p is high), the nonlinear trend filtering generates artifacts, visible as slow changes in the detrended waveform. For that reason a high order MAR filtering cannot be used in each case, especially when a drift is a nonlinear function.

2.3. Polynomial fitting

Polynomial detrending is another high-pass filtering method [1]. This procedure fits the analyzed signal to a polynomial of given order and then subtracts this polynomial from the analyzed signal. If order of the polynomial is equal to one, we are limited to linear detrending – the simplest case.

There is a big variety of commonly used polynomial fitting methods, but the most popular is a method based on least squares criterion. It is also worth to notice that some of methods, especially fitting to higher order polynomials, might not give correct results. We should underline that polynomial detrending can be repeated a few times, even with consecutively increased degree of polynomial until trend component is removed efficiently.

2.4. Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) is a method of signal decomposition without leaving time domain. This is in contrast to other methods, like Fourier transform or wavelet transform. It is suitable for nonlinear and nonstationary trend components [7–10], so it pretends to be efficient method for noise signals analysis.

Time record is decomposed into the finite additive oscillatory components called Intrinsic Mode Function (IMF). There are two conditions which must be satisfied by each IMF:

1. number of extremes and number of zero crossing of the signal must be equal or differ at most by one,
2. the average value of the envelope interpolating local maxima and the envelope interpolating local minima is zero.

In the figure 2 a block diagram of EMD algorithm is presented. In first step local extremes of input signal $x(t)$ should be determined. Based on the determined values, we can interpolate (using spline function) from local maxima an upper envelope $e_u(t)$ and from local minima a lower envelope $e_l(t)$. In the next step we calculate local mean value $m_n(t)$ and determine function $h_n(t)$ which is a first candidate for an IMF component. If $h_n(t)$ satisfies the conditions 1 and 2 it is an IMF component. If not, we take it as a signal being analyzed and repeat previous steps until it satisfies both conditions. When we obtain an IMF component, we subtract it from the original signal and obtain a residual signal $r_n(t)$. Treating now $r_n(t)$ as an analyzed signal and repeating n times these operations, the consecutive IMF components can be determined. Decomposition stops when $r_n(t)$ is a monotonic or a constant function, which means that it is not possible to extract more IMF components (due to the conditions 1 and 2). The exemplary results of EMD decomposition for fast fluctuating process $y(t)$ superimposed on quadratic trend are shown in the figure 3.

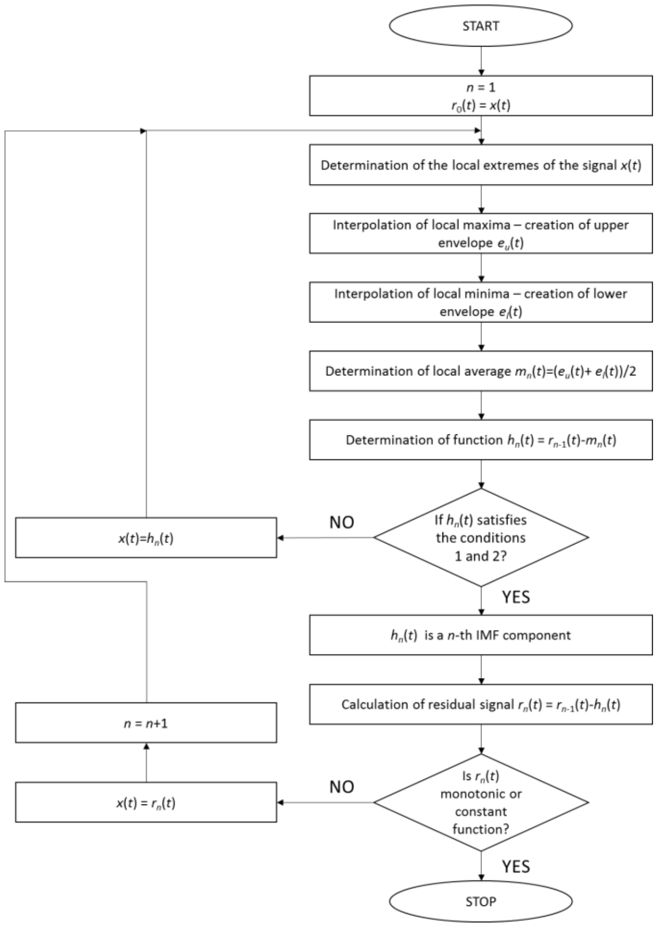

Fig. 2. Block diagram of Empirical Mode Decomposition method

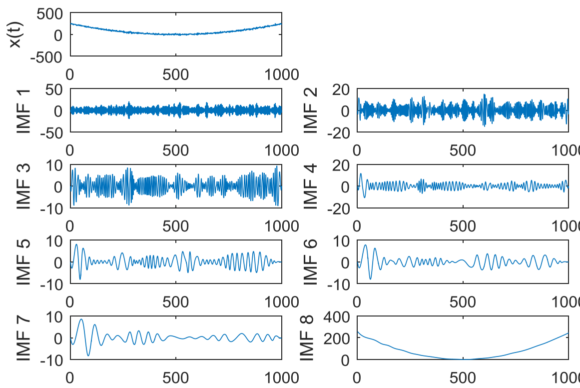

Fig. 3. Exemplary results of EMD decomposition

When an input signal $x(t)$ consists of slowly varying trend superimposed to a relatively fast fluctuating component $y(t)$, which can be e.g. a recorded noise, the trend is expected to be captured by the IMF's of large indexes. The above presented EMD decomposition algorithm can be used to detrended signal $y(t)$ estimation [9]. For this purpose we can use following formula:

$$\hat{y}_D(t) = \sum_{n=1}^{D} h_n(t) \tag{4}$$

where: $\hat{y}_D(t)$ – estimate of the detrended signal $y(t)$ using first D consecutive IMF components, $h_n(t)$ – n-th IMF function, D – the largest IMF index prior contamination by the trend.

A rule of thumb for selection D value is to observe the evolution of the standardized empirical mean of $\hat{y}_d(t)$ as a function of a test value d, and to identify for which $d = D$ it departs significantly from zero.

3. SIMULATION RESULTS

Random component $y(t)$ was simulated in Matlab environment using *randn*() function. Several trend components were simulated and added to $y(t)$ signal to obtain a few signals which were investigated to assess effectiveness of detrending methods. Equations for the investigated signals are presented in Table 1.

Table 1. A set of all investigated signals (x is independent variable)

No.	Equation	Trend component
1	$y(t)+30$	Constant value
2	$y(t)+0.2x+3$	Linear function
3	$y(t)+30\sin(4\pi x/1000)+30$	Sinusoidal function
4	$y(t)+200\exp(-x/100)$	Exponential function
5	$y(t)+0.001(x-500)^2$	Quadratic function
6	$y(t)+200\exp(-x/100)+$ $+10^{-7}(x+800)(x-1200)(x-1200)$	Sum of polynomial and exponential

Probability distribution (histogram – figure 4) and power spectral density (figure 5) were estimated for signal $y(t)$ and then used for testing quality of detrending by means of it similarity, by random error estimation, to those parameters established for the detrended signals.

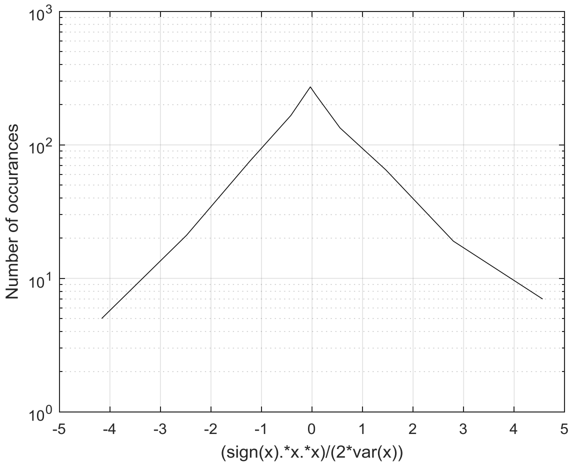

Fig. 4. Histogram of random process $y(t)$ having normal distribution

In the figure 6 we can compare histograms of $y(t)$ signal, $y(t)$ signal combined with sinusoidal trend and detrended signal $\hat{y}(t)$ using three different methods: MAR, polynomial fitting and EMD. It is easy to determine that all histograms for the detrended signals are almost the same as a histogram of original signal $y(t)$, while histogram for original signal superimposed on sinusoidal trend is significantly different.

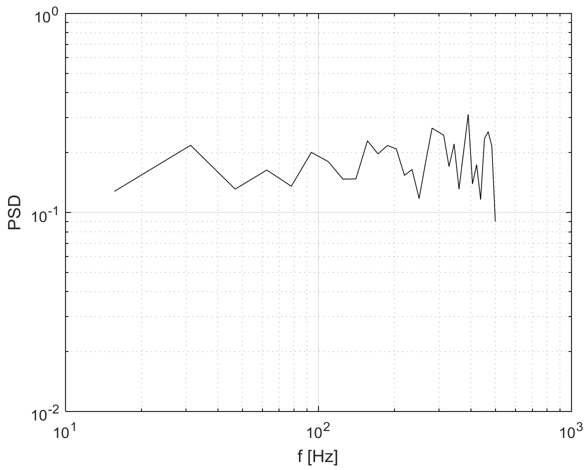

Fig. 5. Power spectral density (sampling frequency was chosen arbitrary) of random process $y(t)$

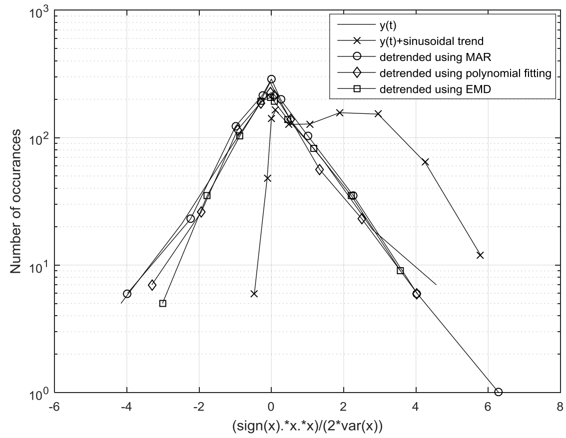

Fig. 6. Histograms of detrended signals by selected methods

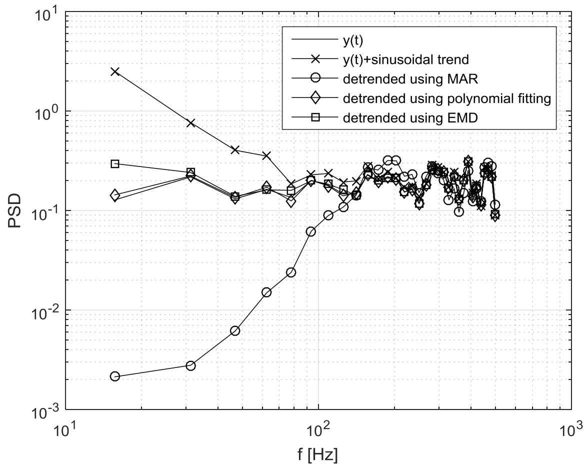

Fig. 7. Power spectral densities of detrended signals by selected methods

Figure 7 compares power spectral densities of $y(t)$ signal, $y(t)$ signal combined with exponential trend and detrended signal $\hat{y}(t)$ using three different methods: MAR, polynomial fitting and EMD. We can notice that power spectral densities of original signal and detrended using polynomial fitting and EMD are almost the same. It is obvious that for a signal with trend component the curve shape is different than in the case of original signal. Spectrum of the signal detrended by MAR method has changed significantly because it should be similar to an applied filter transfer function given by equation (3).

Table 2 includes results of detrending using three mentioned methods and six exemplary signals under investigation (Table 1). Accuracy of trend removal is estimated by root mean squared error (RMSE) between original signal $y(t)$ and detrended signal $\hat{y}(t)$ for all considered methods. We can conclude that efficiency of trend removal depends on type of trend component. Polynomial fitting is very efficient for almost all considered signals except of a signal with sinusoidal trend. The MAR filtering has a constant and highest RMSE due to its low complexity. The EMD shows low RMSE for all considered signals.

Table 2. Accuracy of all investigated methods using exemplary signals

Signal number	Accuracy (RMSE)		
	MAR	Polynomial fitting	EMD
1	3.5	0.3	-
2	3.5	0.6	2.0
3	3.5	6.5	2.1
4	3.5	1.6	2.3
5	3.5	0.8	2.4
6	3.5	0.9	2.5

4. CONCLUSIONS

We showed that power spectral density and probability distribution (histogram) can evaluate quality of trend removal from the recorded random data. Three trend removal methods were investigated: MAR, polynomial fitting and EMD. Efficiency and computational complexity significantly depends on type of the removed trend. The MAR filtering is the simplest method, but its efficiency is rather poor. Good results were observed for the EMD method.

There are numerous applications where the presented methods could be applied. Our considerations were inspired by applying detrending of discharging current in supercapacitors. Another important issue are Raman spectra slowly varying due to bleaching in biological objects.

5. BIBLIOGRAPHY

1. Bertocci U., Huet F., Nogueira R. P., Rousseau P.: Drift Removal Procedures in the Analysis of Electrochemical Noise, Corrosion, no. 4, vol. 58, 2002, p. 337–347.
2. Denholm-Price J. C. W., Rees J. M.: A Practical Example of Low-Frequency Trend Removal, Boundary-Layer Meteorol., no. 1, vol. 86, 1998, p. 181–187.
3. Homborg A. M., Tinga T., Zhang X.,. van Westing E. P. M, Oonincx P. J., de Wit J. H. W., Mol J. M. C.: Time–frequency methods for trend removal in electrochemical noise data, Electrochim. Acta, vol. 70, 2012, p. 199–209.
4. Infield D. G., Hill D. C.: Optimal smoothing for trend removal in short term electricity demand forecasting, IEEE Trans. Power Syst., no. 3, vol. 13, 1998, p. 1115–1120.
5. Mansfeld F., Sun Z., Hsu C., Nagiub A.: Concerning trend removal in electrochemical noise measurements, Corros. Sci., no. 2, vol. 43, 2001, p. 341–352.
6. Vamoş C.: Automatic algorithm for monotone trend removal, Phys. Rev. E, no. 3, vol. 75, 2007, p. 36705.
7. Flandrin P., Rilling G., Goncalves P.: Empirical Mode Decomposition as a Filter Bank, IEEE Signal Process. Lett., no. 2, vol. 11, 2004, pp. 112–114.
8. Moghtaderi A., Flandrin P., Borgnat P.: Trend filtering via empirical mode decompositions, Comput. Stat. Data, vol. 58, 2013, p. 114-126.
9. Flandrin P., Es P., Rilling G.: Detrending and denoising with empirical mode decompositions, Signal Processing Conference, IEEE 2004 12th European, p. 1581-1584.
10. Rilling G., Flandrin P., Gon P., Lyon D.: On Empirical Mode Decomposition and Its Algorithms, IEEEEURASIP Work. Nonlinear Signal Image Process. NSIP, vol. 3, 2003, p. 8–11.

ACKNOWLEDGEMENT

This work was supported by the National Science Center, Poland, the grant decision: DEC-2014/15/B/ST4/04957 "Mechanizm procesów ładowania/wyładowania na granicy faz elektroda/elektrolit w superkondensatorach".

ANALIZA SKUTECZNOŚCI I ZŁOŻONOŚCI OBLICZENIOWEJ METOD USUWANIA SKŁADOWEJ TRENDU Z DANYCH POMIAROWYCH

W pracy przedstawiono sposób przetwarzania danych pomiarowych w celu usunięcia wolnozmiennej składowej trendu występującego w rejestrowanych przebiegach. Porównano kilka często stosowanych w tym celu metod pod względem ich złożoności obliczeniowej oraz skuteczności w usuwaniu trendu. Pokazano wpływ tych procedur na rozkład prawdopodobieństwa wartości chwilowych oraz przebieg gęstości widmowej mocy. W ogólności operację usuwania trendu możemy traktować jako filtrację górnoprzepustową danych pomiarowych. W celu usunięcia trendu można użyć filtru górnoprzepustowego (analogowego lub cyfrowego) już na etapie akwizycji danych pomiarowych. Jednakże często mamy do czynienia z danymi, w których składowa trendu jest potrzebna do przeprowadzania innych analiz i nie może być usunięta na etapie rejestracji danych pomiarowych. Ponadto, może mieć charakter niestacjonarny i metody filtracji górnoprzepustowej nie będą skuteczne. W takich przypadkach należy rozważyć inne, często bardziej zaawansowane metody. Skuteczność i złożoność obliczeniowa takich metod zależy istotnie od charakteru usuwanego trendu. W pracy opisano procedurę usuwania średniej kroczącej (ang. Moving Average Removal – MAR), metody o niskiej złożoności obliczeniowej, ale dającej zadowalające rezultaty w dużej liczbie potencjalnych zastosowań. Rozważono usuwanie trendu przez dopasowanie wielomianem odpowiedniego stopnia do analizowanych danych pomiarowy. Procedura ta może być powtarzana kilkukrotnie, nawet ze zwiększaniem stopnia wielomianu przy każdym z kroków, aż do uzyskania przebiegu, w którym usunięto składową trendu. Część pracy poświęcono prezentacji bardziej złożonych obliczeniowo metod, które zostały rozwinięte dopiero w ostatnich latach i wymagają znacznie bardziej intensywnych obliczeń.

Słowa kluczowe: usuwanie trendu, filtracja górnoprzepustowa, empiryczna metoda dekompozycji.